

Module 6 Lab Activity: Correlation, Partial Correlation & SLR
PSY 652 Research Methods
Oct 7, 2020

Dataset Description: A researcher was interested in assessing the relationship between caloric deficit and pounds lost over the course of a 1 month weight loss program designed for men age 40 to 50 in CO. 100 obese men were randomly selected to participate. The researcher measured the following:

Variables:

1. **lbslost** = pounds lost after one month on the weight loss program
2. **caldef** = caloric deficit over the course of the program (expressed in 1000 calories). A caloric deficit is a state in which you are burning more calories than you eat.
3. **selfeff** = self-efficacy for weight loss at the start of the program
4. **id** = subject ID number

Activity Instructions:

1. Create a new R notebook from your project file and name it “wtloss_notebook”
2. Download the “wtloss_parta.csv” and “wtloss_partb.csv” datasets from the Module 6 Lab module.
3. Create a new R chunk with a first level header: “Load Libraries”
 - a. load the ppcor, psych, apaTables, and tidyverse packages in this R chunk (you may need to install some of these)
4. Create a new R chunk with a first level header: “Import Data”
 - a. read in the “wtloss_parta.csv” and assign it to an object names “wtlossa”
 - b. read in the “wtloss_partb.csv” and assign it to an object names “wtlossb”
5. Create a new R chunk with a first level header: “Merge (Join) Data”
 - a. Merge the wtlossa and wtlossb dataframes together by subject ID number and assign this new dataframe to an object called “wtloss”. Confirm that this merge worked by clicking on the wtloss dataframe in your global environment (this should include 100 observations of 4 variables).
 - i. Hint: `left_join(data_a, data_b, by = “id_variable”)`
6. Write a first level header: “Describe Data”
 - a. Create a new R chunk with the second level header: Calculate descriptives for the wtloss dataframe using the describe() function from the psych package.

7. Write a first level header: “Plot relationships between variables”
 - a. Using `ggplot()`, create a scatterplot of the relationship between: *lbslost* & *caldef*. Make sure *lbslost* is on your Y-axis. Add the best fit linear line `geom_` layer (`geom_smooth`).

Hint: `ggplot(dataset, aes(x = var1, y = var2)) +
geom_point() +
geom_smooth(method = “lm”)`
8. Write a first level header: “Correlations”
9. Write a second level header: “Pearson’s Correlation Matrices”
 - a. Create a new R chunk with the third level header: “Via base R’s `cor` function.” Create a correlation matrix for all variables in the *wtloss* dataframe using the `cor` function.
 - b. Create a new R chunk with the third level header: “Via `apaTables`’ `apa.cor.table` function.” Create a correlation matrix for all variables in the *wtloss* dataframe using the `apa.cor.table` function.

Hint: `apa.cor.table(dataframe_name, “file_name.doc”,
show.conf.interval=TRUE)`

Note: The `apa.cor.table` function creates a .doc form of this correlation matrix, which will be saved in your folder for this R notebook. Open this document in your project folder to make sure it worked. This package can be a big time-saver when preparing tables for publication!
 - c. In the white space below this chunk, describe each correlation. What direction are the correlations? Are they strong or weak correlations? (Note: you don’t need to interpret subject id number here)
10. Write a second level header: “Get Partial Correlation”
 - a. In a new R-chunk, use the `pcor.test()` function to get the partial correlation between *lbslost* and *caldef*, while controlling for *selfeff*.
 - i. Hint: `pcor.test(x = data$var1, y = data$var2, z = data$var3)`
 - b. In your own words, interpret the estimate and p-value of the output.
11. Write a first level header: “Get Semipartial Correlation”
 - a. In a new R-chunk, use the `spcor.test` function to get the semipartial correlation between *lbslost* and *caldef*, while controlling for *selfeff*.
 - i. Hint: `spcor.test(x = data$var1, y = data$var2, z = data$var3)`
 - b. In your own words, interpret the estimate and p-value of the output.
12. Create a first level header: “Fit a Simple Linear Regression (SLR) Model”
 - a. In a new R chunk, create a linear model in which *caldef* predicts *lbslost*. Save this model to an object called “*mod1*”

Hint: `model_name <- lm(y ~ x, data = dataframe_name)`
 - b. Write a second level header: “Display SLR Model Results via `summary()`”

- i. Insert a new R chunk: Use the `summary()` function to display the model's results.
 - c. In the white space below, interpret your model output, including:
 - i. the estimate for the intercept
 - ii. the estimate for *caldef*
 - iii. the *p* value for the *caldef* estimate
 - iv. the model R^2
13. Once you've completed all of these steps, Restart R and Run All Chunks, and then preview your notebook. Save your notebook as both a .Rmd and an html file and exit RStudio.
14. Upload the html version of your notebook to the assignment called "Upload Module 6 Lab Activity Here" on the course Canvas page.